

Performance Measurement in Healthcare:
Part II – State of the Science Findings
by Stage of the Performance
Measurement Process

La mesure du rendement dans les soins de santé :
Partie II – Résultats de l'examen de l'état
de la science, par étape du processus de
mesure du rendement



by CAROL E. ADAIR, MSC, PHD
*Associate Professor, Departments of Community Health Sciences and Psychiatry
University of Calgary, Calgary, AB*

ELIZABETH SIMPSON, BA, MSC
*Health Research Consultant
Red Deer, AB*

ANN L. CASEBEER, MPA, PHD
*Associate Professor, Department of Community Health Sciences
Associate Director, Centre for Health and Policy Studies
University of Calgary, Calgary, AB*

Performance Measurement in Healthcare: Part II

JUDITH M. BIRDSELL, BSCN, MSC, PHD
Principal Consultant, ON Management Ltd.
Adjunct Associate Professor, Haskayne School of Business
University of Calgary, Calgary, AB

KATHARINE A. HAYDEN, MLIS, MSC, PHD
Associate Librarian, Information Resources
University of Calgary, Calgary, AB

STEVEN LEWIS, BA, MA
President, Access Consulting Ltd.
Adjunct Professor, Department of Community Health Sciences
University of Calgary, Calgary, AB

Abstract

Objective: This paper summarizes findings of a comprehensive, systematic review of the peer-reviewed and grey literature on performance measurement according to each stage of the performance measurement process – conceptualization, selection and development, data collection, and reporting and use. It also outlines implications for practice.

Methods: Six hundred sixty-four articles about organizational performance measurement from the health and business literature were reviewed after systematic searches of the literature, multi-rater relevancy ratings, citation checks and expert author nominations. Key themes were extracted and summarized from the most highly rated papers for each performance measurement stage.

Results: Despite a virtually universal consensus on the potential benefits of performance measurement, little evidence currently exists to guide practice in healthcare. Issues in conceptualizing systems include strategic alignment and scope. There are debates on the criteria for selecting measures and on the types and quality of measures. Implementation of data collection and analysis systems is complex and costly, and challenges persist in reporting results, preventing unintended effects and putting findings for improvement into action.

Conclusion: There is a need for further development and refinement of performance measures and measurement systems, with a particular focus on strategies to ensure that performance measurement leads to healthcare improvement.

Résumé

Objectif : Ce document résume les résultats d'un examen détaillé et systématique de la littérature grise et des publications évaluées par les pairs sur la mesure du rendement pour chaque étape du processus – conceptualisation, sélection et développement, collecte de données, présentation des résultats et utilisation. Il présente aussi des répercussions sur la pratique.

Méthodes : Après avoir effectué des recherches systématiques dans la littérature, demandé à des évaluateurs multiples de déterminer la pertinence des documents repérés, vérifié les citations et désigné les auteurs experts, 664 articles sur la mesure du rendement organisationnel provenant de publications des domaines de la santé et des affaires ont été examinés. On a dégagé puis résumé des thèmes clés à partir des documents ayant reçu la plus haute cote pour chaque étape de la mesure du rendement.

Résultats : Malgré un consensus quasi universel sur les avantages potentiels de la mesure du rendement, il existe actuellement peu de preuves pour guider la pratique dans les soins de santé. Les problèmes de conceptualisation des systèmes comprennent, entre autres, l'alignement stratégique et la portée. On ne s'entend pas sur les critères à utiliser pour sélectionner les mesures et sur les types et la qualité de ces dernières. La mise en place des systèmes de collecte et d'analyse de données est complexe et coûteuse, et il y a encore des défis à relever dans la présentation des résultats, la prévention des effets non prévus et la transformation des résultats en des mesures concrètes.

Conclusion : Il faut développer et peaufiner davantage les mesures du rendement et les systèmes connexes, en mettant un accent particulier sur les stratégies pouvant garantir que la mesure du rendement mènera à des améliorations dans les soins de santé.

THE PURPOSE OF OUR REVIEW WAS TO SUMMARIZE THE CURRENT BUSINESS and healthcare literature on performance measurement (PM) systems and to make recommendations for research and practice. Details of methods are provided in Part I (*Healthcare Policy*, 1.4). This second paper reports in greater depth on themes and issues extracted from the peer-reviewed and grey literature in relation to stages of the PM process.

The PM Process

The PM literature lacks consensus on concepts and definitions. However, the PM process is typically described as having approximately four stages (Nadzam and Nelson 1997; Nutley and Smith 1998; Bourne et al. 2000; Ibrahim 2001; Smith and Goddard 2002), although many authors caution that the process is more dynamic and less linear than a simple set of stages implies. The stages are (a) conceptualization, (b)

selection and/or development of measures, (c) data collection and processing and (d) reporting and using results.

Conceptualization

Two major issues on conceptualization of PM systems are prominent in the literature: aligning with organizational *strategic direction* and determining the appropriate *scope for the system*.

STRATEGY

There is increasing emphasis on aligning PM activities with the strategic direction of the organization, and a general sentiment in both business and health that such alignment is rare in practice. However, maintaining a strategic focus is acknowledged to be more difficult in healthcare than in business for several reasons.

First, organizational goals are often difficult to operationalize in healthcare because of the complexity of treatments, settings and patient groups (Baker and Pink 1995). Public service organizations have broader goals (including societal goals) and “a more complex pattern of accountability than the corporate financial statement” (Smith 1993: 137). The dual management model (professional and administrative) and the interrelationships among multiple internal and external stakeholders (Kleinpell 1997; Lemieux-Charles et al. 2002), each with its particular interest in setting the PM agenda (Nadzam and Nelson 1997; Collopy 1998), create greater complexity. In health services the policy environment is very fluid (Smith and Goddard 2002), perhaps more so than in business environments.

Second, causal links between service and health outcomes are very difficult to specify for both medical and public health interventions, owing to the limits of evidence in medicine and the reality that healthcare is only one of several predictors of health status (Williams et al. 1992; Handler et al. 2001; Leggat et al. 1998).

Third, “customer” dynamics are less straightforward in healthcare than in the purchase of a commercial product or service (Newhouse 2002). People seek care out of necessity, not desire. The provider often has a local monopoly on a given service, limiting both comparators for judgments about performance and opportunities to seek alternatives (Smith 1993). An important commercial goal is repeat business, while in healthcare it is often viewed as an unfortunate necessity because a definitive cure is unattainable. The consumer is also typically less knowledgeable about the service content than in commercial transactions (Jennings and Staggers 1999) and is often vulnerable by virtue of being ill and possibly afraid when seeking care. These realities complicate the patient satisfaction and perceived care quality domains of PM (Jennings and Staggers 1999). The message about the task of strategic conceptualiza-

tion of a PM system is clear in both sets of literature: “what gets measured gets delivered,” and there are undesirable consequences for organizations, from a strategic point of view, that collect the wrong measures (Voelker et al. 2001).

SCOPE

The second major issue in conceptualization of PM systems in both literatures is determining the appropriate system scope. Scope decisions apply to three dimensions: vertical (level of the healthcare organization or system), horizontal (breadth across the continuum of care or business units) and longitudinal (temporal) (Collopy 1998). In business there is a trend towards involving all levels of the organization in a common vision that can be reinforced by the PM system itself (Neely et al. 1995; Epstein and Manzoni 1998; Lockamy 1998; Legnini et al. 2000). “One of the major problems with conventional PM is the ease with which organizational wholes are carved up, and their interactions with their environments cease to be of interest as management functions devise measures (and associated targets) for their own territory. This reductionism is associated with some of the problems identified by managers when they seek to improve performance” (Holloway 2001: 173).

Healthcare PM activities are also highly fragmented, verified by the sheer number of single-level or single-service systems described in the literature. Single-level focus creates debates about the value of one over the other: some charge that the patient level is often not addressed in system-level approaches (e.g., Greenhalgh et al. 1996), while others express the opposite concern (e.g., Barrell 2000). Many call for greater consolidation through overarching goals and greater consensus and coordination (Eddy 1998; Kizer 2001), and increasingly multi-level systems are being conceptualized (e.g., Moscovice et al. 1995; Luttmann 1998; Evans et al. 2001; Handler et al. 2001). Even so, Nutley and Smith (1998: 53) contend that “calls for a top to bottom PM architecture have largely been ignored.” Others caution that the PM for high-level management and accountability differs from that needed for daily operations (McLoughlin et al. 2001; Voelker et al. 2001).

The horizontal scope of systems is also debated. The business literature reports a few companies attempting to establish measures that capture relevant information across company boundaries (such as with supplier networks), but acknowledges this to be very difficult (Fawcett and Cooper 1998). The roots of healthcare PM are clearly in acute care, and hospital-bounded approaches dominate. Separate PM systems are under development and are testing for other components such as public health (Corso et al. 2000; Handler et al. 2001; Kates et al. 2001), but our review found no systems spanning acute and community care. DeRosario (1999: 38) notes that “to catch the next wave of performance change, we need to begin measuring activities that occur between healthcare sectors,” and others concur (Hall 1996; Kizer 2001). A PM sys-

tem should match the service delivery model, and it is likely that broader PM systems will emerge with the trend towards regionalized, integrated health services in many jurisdictions. With respect to the temporal dimension, a few authors suggest that PM systems need to address and measure the process of care over time for an individual (Bishop and Pelletier 2001).

Measures selection or development

Many authors stress that, according to measurement theory, measures themselves are just a reflection of reality. In addition, the choice of what to measure among the many options is an imprecise process (van Peurse et al. 1995), reflecting a system of values and social goals (Sheldon 1998). Ibrahim (2001: 431) writes that “performance indicators are inherently controversial” because they require a judgment about what constitutes quality.

FRAMEWORKS

After general conceptualization, the next task in PM is to select or develop measures. Optimally, a framework ensures balance across strategic improvement areas and guides the measurement process. An ideal framework describes domains (measure groupings) and dimensions (e.g., organizational levels), but most frameworks reviewed are simply a list of indicators and/or domains (e.g., Lied 1999). More complex frameworks also include one or more dimensions such as level of the healthcare system (McEwan and Goldner 2000) or stakeholder perspective (Nadzam and Nelson 1997; Kizer 2001; McIntyre et al. 2001). We found little consistency in the combinations of 21 domains used in 17 major health PM frameworks reviewed (Adair et al. 2003).

We identified eight business frameworks that included both non-financial and financial measures (Lebas 1995; Neely et al. 1995; Kaplan and Norton 1996, 2001; Epstein and Manzoni 1998; Kueng and Krahn 1999; Kueng 2000; Kanji and Moura 2002) – called multi-dimensional or portfolio approaches – that are tabulated in the full report (Adair et al. 2003). Neely et al. (2000) and Kueng (2000) provide noteworthy reviews of business approaches. The most popular framework in business is the Balanced Scorecard (BSC), which has also been applied in healthcare. Some other approaches to the management of quality in the business literature are noteworthy because of their recent diffusion into healthcare and their close relationship with PM. First are the quality award programs, including the Malcolm Baldrige National Quality Award, the European Foundation for Quality Management’s Business Excellence Model (Neely et al. 1995; Kueng and Krahn 1999; DeBaylo 1999) and many spin-off quality award programs. One widely adopted program, Hoshin Kanri, that developed in Japan in the 1960s and has been disseminated widely is noteworthy

for having extensive coverage in the popular press worldwide but virtually none in the western research literature (Tennant and Roberts 2000). The BSC and other portfolio approaches have evolved towards the selection of more forward-looking, strategy-focused measures, but many criticisms of these early-stage approaches persist (Kueng and Krahn 1999; Mooraj et al. 1999; Kueng 2000; Baughan et al. 2002; Brignall 2002; Morgan and Braganza 2002) that parallel the healthcare PM literature.

ISSUES IN CHOOSING MEASURES

Several predominant themes relate to measures selection, including the sheer growth in numbers of measures and systems, as well as issues related to the types of measures and their limitations.

In recent years, measures (both indicators and comprehensive instruments) have become so numerous that it would be nearly impossible to catalogue them completely (Nutley and Smith 1998; Sheldon 1998). The national indicator library of the Joint Commission on Accreditation of Healthcare Organizations (JCAHO) is believed to have more than 1,000 measures, and the database of the Agency for Healthcare Research and Quality (AHRQ) contained more than 1,197 in 53 sets by 1995 (AHRQ 2002). Unless indicators are commonly defined, comparative reporting is difficult, if not impossible. The development of measures databases is a welcome sign that this duplication of effort may be waning (e.g., Jennings and Staggers 1999; Hermann et al. 2000). Collaborative efforts to standardize measures are another promising development (Braun and Zibrat 1996; Leggat et al. 1998).

Guidelines or criteria for indicator selection are numerous in both literatures and, again, there is little consistency across sets. Table 1 lists criteria catalogued and synthesized conceptually from health literature papers that are cited in the full report but are too numerous to cite here (Adair et al. 2003). They represent suggested, rather than tested, criteria. The more recent literature puts greater emphasis on the importance of choosing indicators that are meaningful, strategic and evidence-based.

Financial indicators are still used as part of health PM systems (e.g., cost per weighted case), but as in business, non-financial indicators have taken centre stage. In discussing BSC applications in health, Voelker et al. (2001) claim that a primary focus on financial measures may actually hinder organizational growth and success. In healthcare, financial measures are notoriously difficult to action because most costs are not variable and there is little flexibility in hiring and firing staff (Brookfield 1992). Because of the complex and multifaceted purposes of healthcare, focusing too heavily on financial measures may diminish prospects for overall improvement. Most PM systems in health continue to collect traditional input/output measures such as service utilization (e.g., bed occupancy, surgery facility use, length of stay and numbers of discharges and admissions), despite repeated commentary that they are poor indi-

TABLE 1. **Criteria for performance measures selection**

CRITERION	DESCRIPTION
Evidence-based	There are valid and reliable operational definitions for the measure that have been demonstrated through rigorous research
Strategic	The measure directs attention towards the ultimate change desired
Important	The measure addresses an important or serious health or health services problem (usually defined as health burden or cost) such that there will be sufficient impact from collection and service improvement initiatives
Attributable	Causal links between the measure, service improvements and health outcomes are known
Actionable	The measure addresses a service area that can benefit from improvement
Feasible	Data collection, reporting and follow-through are cost-effective (potential benefits outweigh costs) and there is reasonable technical capacity for collection and analysis, including risk adjustment of compared measures
Relevant and meaningful	The measure is relevant to most stakeholders, including policy makers, managers, clinicians and the public
Understandable	The measure is understandable to a non-technical audience (often just a communication issue)
Balanced	The set of measures is balanced across types of treatments, treatment settings, major health problems, age groups, special populations and levels of the healthcare system. The set is balanced across short- and long-term measures, and balance and appropriateness are considered across process- and outcome-type measures
Responsive	The measure is sensitive to change over time
Robustness	Potential adverse effects of the measure can be mitigated, and vulnerability to gaming is minimal
Non-ambiguous	The measure is clear in terms of which direction for service change is desirable

cators of performance (Mark et al. 1997; Nutley and Smith 1998). Mortality remains the predominant traditional outcome measure, with the distinct disadvantage that it reflects a rare and end-stage event relative to the total volume of healthcare provided. In a Canadian study of existing indicators reported in 2000, Lemieux-Charles et al. (2000: 52) observed that “indicators measuring integration, coordination and continuity of care, as well as responding to population health needs, were rarely used. These types of measures are critical as we redesign our service delivery systems to address population needs.” Klazinga et al. (2001) consider the ultimate performance measures to be those reflecting overall population health.

Similarly, others express concern about “opportunistic systems” that emphasize readily available measures at the expense of newer, more important and meaningful measures (West 1996; Elkan and Robinson 1998; Nutley and Smith 1998; Smith and Goddard 2002). Shaw (1997: 217) characterizes this as the “spectre of convenience” and asks, “should measures be based on existing available data as ad hoc criteria for achievement, or should health service policy targets first be identified and data then captured specifically to measure their achievement?” A dynamic tension exists between the need for locally meaningful and strategic measures and the benefits of selecting and using standardized measures that enable meaningful comparison.

The business literature also underscores the point that the choice about what not to measure is as important as what to measure, since “things that are measured are considered important while the things not measured are generally considered of less importance” (Waggoner et al. 1999: 54). This literature also notes that once collected, measures are rarely deleted, even if they are obsolete (Neely et al. 2000). Given limited resources, each measure chosen represents an opportunity cost.

The component literatures reveal an important parallel debate about process versus outcomes measures (e.g., Evans et al. 2001; Rubin et al. 2001; Mannion and Davies 2002). The business literature uses other terms, e.g., “a debate on whether performance indicators should be focused on procedures (activities) or on results (output)” (Kueng 2000: 77), but the concepts are identical. Despite some arguments that process measures are more practical, most writers consider them complementary to outcomes or results (e.g., Baker 1995), and all should be chosen to fulfill the specific measurement objective (Wynia et al. 1996).

There are widespread concerns about the paucity of validation work. Eddy (1998: 7) describes current measures as “blunt, expensive, incomplete, and distorting.” There is strong consensus that measures must be evidence-based. Gross et al. (2000) evaluated coronary bypass mortality-related indicators across 24 hospitals and concluded that indicator definitions significantly affected computed rates and changed relative standings. “There are no generally agreed-on external criteria for validity of indicators” (Gross et al. 2000: 210).

Data collection and analysis

Both component literatures strongly note the unanticipated cost and complexity of PM systems. The business literature describes data collection and analysis as “complex, frustrating, difficult, challenging, important, abused and misused” (Lebas 1995: 23). Costs rise because of the high level of technical and managerial expertise required, new information technology and ongoing maintenance. Some also attribute costs (monetary and strategic) to measuring too many different things. “Measuring something makes it important and therefore motivates people. Measuring everything means nothing is important and therefore de-motivates” (Johnston and Fitzgerald 2001: 183). Kueng (2000) identifies success factors in the data collection stage as a parsimonious set of generally accepted indicators, automation and personal involvement of staff and management.

In healthcare, many organizations have lacked the capacity to implement effective systems, and failed attempts are abundant. They generally underestimate the scope and complexity of the infrastructure required to manage healthcare adequately and, by implication, the measurement of its performance (McIntyre et al. 2001). Voelker et al. (2001) and Braun and Zibrat (1996) attribute system failures at this stage to staff and management turnover, technical problems with information systems, budget constraints and competing priorities. Kates et al. (2001) express concern about mandating PM systems in public service organizations without guidance in their implementation and use. Both literatures express concerns about the cost–benefit relation of PM initiatives.

Other issues related to data collection include data sources and quality. Administrative data have long been considered a rich source for PM if properly “mined,” and researchers in particular have produced notable examples of their creative and rigorous use (e.g., Brownell et al. 2001). But many now suggest that the value of secondary data has been overstated, at least as typically formatted (Bishop and Pelletier 2001; McLoughlin et al. 2001). Problems cited include poor reflection of performance, lack of data elements for sensitive diagnosis and risk adjustment, lack of availability and stability of data at smaller levels of aggregation and generally poor quality (Kelman and Smith 2000; Brown 2002). Many writers bemoan the effort devoted to the analysis of retrospective or secondary data at the expense of the collection of more relevant data (Sheldon 1994; Stryer et al. 2000; Voelker et al. 2001). In the more general context of effectiveness research, after 10 years of experience with secondary data, AHRQ’s Patient Outcome Research Team (PORT) investigators are also calling for more prospective and real-time data (Stryer et al. 2000).

Many advocate for routine prospective data collection, fully integrated with clinical practice, that can be used for the delivery of care as well as rolled up for management use (McLoughlin et al. 2001). Concerns remain about the diversion of clinician time from patient care to data recording tasks (Naylor 1999). Ullman et al. (1996: 361) suggest that research-based, standardized measures are “too unwieldy and time

consuming to mesh well with the practice ecology.” Several hybrid approaches are proposed (e.g., Schneider et al. 1999; Brook et al. 2000; Hoelzer et al. 2001), and many commentators still consider the electronic health record, with the appropriate data for PM thoughtfully built in and integrated with more general operational data, to be the best solution in the long run (Aller 1996; Slater 1997).

The literature is replete with concerns about PM data quality. These include issues of missing data, reliability, validity, accuracy, precision, statistical and clinical significance and timeliness (Kleinpell 1997; Mark et al. 1997; Shaw 1997; Collopy 1998; Jencks 2000; Roper and Mays 2000; Pink et al. 2001). McKee and James (1997) provide an excellent review of data quality issues that arise when comparing outcomes data across systems that use different diagnostic and severity adjustment schemes, and report error rates as high as 20% to 40%. Many cite the need for consistent definitions and processes and data quality checks (Shaw 1997; Nutley and Smith 1998) and for the transparent reporting of data collection issues that underlie the reported measures (Pink et al. 2001). Pink et al. (2001) consider expert involvement of both researchers and management as essential.

With respect to methods for analysis, sound statistical methods have long been available but many authors suggest that they usually fall by the wayside in practice (Leggat et al. 1998; Nutley and Smith 1998; Roper and Mays 2000; Smith and Goddard 2002). Adjustment methods are many and varied, and consensus is lacking about the best methods for a given analytic problem (Mant and Hicks 1996; Iezzoni 1997; Shahian et al. 2001; Schneider 2002; Smith and Goddard 2002). Several authors stress that the problem is not so much the methods’ mechanics but the lack of understanding of their limitations and inconsistency in application (Ibrahim 2001; Zaslavsky 2001). An obvious solution is to ensure that adequate analytic expertise is brought to the PM task. Organizational comparisons should disclose all analytic methods and reveal potential sources of bias. As well, a “healthy skepticism about ratings or ranking [should] be maintained” (Schneider 2002: 3). Smith and Goddard (2002) suggest that devising better ways to communicate complex results to non-experts could strengthen the link between research and strategic policy.

Reporting and use

A first general theme on the topic of reporting PM information is practical advice on effective presentation for various audiences, with the emphasis on evidence-based communications. A more prominent and controversial topic is the growing practice of reporting performance information to external stakeholders via report cards. Several authors provide excellent reviews of the issues and evidence related to public release of performance data (Leatherman and McCarthy 1999; Marshall et al. 2000; Hoey et al. 2002). Barrell (2000: 15) expresses the general sentiment on this matter: “There seem

to be basically two schools of thought: those who believe we can't afford to do it, and those who believe we can't afford not to." In a rare and interesting empirical study that examined organizational response to public disclosure of quality data in the United States, McCormick et al. (2002) demonstrated that in a voluntary system, providers with lower-quality scores were four to six times more likely to withdraw from future disclosure than those with higher scores.

We also found a large literature on the issue of using PM to produce improvement. The business literature clearly advocates a strong link between performance *measurement* and performance *management* (Lebas 1995), including the development of causal models between measures, actions taken and subsequent improvement (Lebas 1995; Neely et al. 1995; Neely 1999) through an organizational change process (Kueng 2000). With respect to alignment of incentives for change, Epstein and Manzoni (1998) cite Kerr's folly (rewarding A while hoping for B) as a common practice in many companies, due to an inability to break out of old patterns of reward and recognition, the lack of an overall system view and focusing on the short term.

The health literature addresses three themes on the application of PM information: its use by organizations as a whole, by individual service providers and, externally, by consumers to make care choices. A second theme is how PM is used for both positive change as well as its unintended or adverse effects. A third is the organizational culture in which PM is embedded.

First, on the issue of "actioning" results, Goddard et al. (2000: 99) observe that "most schemes appear to rely on a vague hope that providers will 'do something' in response to the data." The importance of organizations learning how to link the PM results to actions, rather than having the PM system simply keep records, is restated in many ways (Camp and Tweet 1994; Baker and Pink 1995; Collopy 1998; Voelker et al. 2001). The few studies on organizational (Turpin et al. 1996; Leggat et al. 1998; Lemieux-Charles et al. 2000) or individual provider behavioural change (Jencks 2000; Marshall et al. 2000) in response to organizationwide PM suggest that impact is minimal (Barrell 2000; Legnini et al. 2000; Marshall et al. 2000; Schneider 2002). It is likely that in some settings individual managers and clinical leaders have found effective ways to use and apply performance measurement information, just as in some settings quality improvement has been applied effectively – many examples are provided by the Institute for Healthcare Improvement (2002) – but virtually no rigorous studies have described effective broader-level PM practice and elucidated its features.

The more recent healthcare literature includes descriptions of new mechanisms involving financial incentives for performance at the organizational or individual level. These mechanisms go by a variety of labels, including value-based purchasing, quality-based purchasing, performance-based contracting and pay-for-performance. With respect to alignment of financial incentives at the organizational level, there were many reported instances in US healthcare and some in the United Kingdom. A straight-

forward incentive system that simply provides high performers with extra funds and penalizes low performers is criticized as having the potential to flow funds to services serving regions with less health need, if the contributors to poorer performance are environmental and socio-economic rather than actual differences in care (Elkan and Robinson 1998). In a fairly innovative concept for incentive alignment, Ward (2000) describes a scheme for improving performance in NHS trusts. In this scheme, funding is not allocated according to performance ranking; instead, *greater autonomy and spending latitude* are given to higher-ranking organizations (Ward 2000). While financial incentives may seem like common sense, they continue to be controversial and are largely unproven to date (e.g., Giuffrida et al. 2000).

With respect to the potential for adverse effects, the literature contains many examples of (mostly theoretical) adverse effects, which are summarized in Table 2. Goddard et al. (1998, 2000), Smith (2002) and Smith and Goddard (2002) have drawn from the management control literature and written extensively on unintended effects in the public sector and healthcare. They consider that “some of these dysfunctional consequences are the result of the imperfect or incomplete data on which indicators are based, some are due to how the data are used and interpreted, and some are simply intrinsic to any system of PM” (Goddard et al. 1998: 26).

TABLE 2. Unintended or adverse effects of performance measurement

1. Attention can be focused narrowly on improvement of the measure itself, rather than the underlying process
2. Measures can be selected that divert attention and effort away from more important problems, or measures can be focused on the short term at the expense of longer-term issues
3. Measurement may encourage an attitude of seeking simplistic solutions to complex problems
4. Individual managers can use measurement to serve their own agendas rather than the needs or priorities of the whole organization
5. Measures can be “gamed” or distorted
6. Average performance may be considered sufficient, encouraging complacency and discouraging risk-taking
7. Measures can be used to lay blame rather than find solutions
8. Good results are disseminated while poorer results are suppressed
9. Broader performance expectations or standards can dominate local priorities
10. Unrealistic performance targets can lower morale and engender defeatism

Sources: Smith 2002; Smith and Goddard 2002; Goddard et al. 1998, 1999, 2000; van Peursem et al. 1995; Collopy 1998; Elkan and Robinson 1998; Leggat et al. 1998; Proctor and Campbell 1999; McLoughlin et al. 2001.

A third theme in the health literature is the relatively recent acknowledgment that organizational contextual issues are paramount to effective PM use because of the invariably complex health system environments. Smith (1993: 150) suggests that while PM systems are assumed to be neutral reporting devices, in reality they are “operating in a far messier and less well understood organizational context.” Barnsley et al. (1996), Leggat et al. (1998) and others outline the organizational culture issues in PM. Legnini et al. (2000) provide a very detailed set of recommendations for realigning incentives to encourage positive use of PM information, according to organizational context and stakeholder perspective. Table 3 lists other suggestions. A more comprehensive and holistic approach to PM is being promoted (McKee and Sheldon 1998; Smith 2002), and the emergence of new models may be imminent (Viccars 1998; Campbell et al. 2001).

TABLE 3. Suggested solutions

1. Leadership and commitment of senior managers/decision-makers is essential
2. Take a systems approach, including consideration of organizational, contextual issues
3. Focus on positive personal development, including education, supports for role change and realignment of incentives
4. Maintain a positive, constructive, solution-focused orientation, not a blaming approach
5. Consider performance measures as flags for identification of areas for improvement, rather than absolute measures of performance
6. Commit to PM as a long-term endeavour
7. Resource PM appropriately; ensure that the appropriate technical and managerial expertise and adequate funds are available
8. Foster continuous, open communication with emphasis on interpretation of findings, avoiding simplistic explanations
9. Encourage ownership of PM through collaborative, participatory approaches
10. Consider all stakeholders' perspectives
11. Plan for performance management, not just measurement, i.e., ensure that mechanisms are in place to use results

Sources: Greenhalgh et al. 1996; Mant and Hicks 1996; Turpin et al. 1996; Ford et al. 1997; Collopy 1998; Goddard et al. 1998; Leggat et al. 1998; Nutley and Smith 1998; Bodenheimer 1999; Proctor and Campbell 1999; Gross et al. 2000; Voelker et al. 2001; Weinberg 2001; Zairi and Jarrar 2001; Inamdar et al. 2002; Jarvi et al. 2002; Mannion and Davies 2002.

Summary and Implications for Practice

The literature reviewed on PM reveals several points of consensus as well as divergence, as summarized in Table 4. Overall, no author advocated abandonment of PM, but most recommended moving forward with more awareness of the pitfalls and mak-

ing informed choices (Smith 1993; van Peurse et al. 1995; Shaw 1997; Eddy 1998; Sennett 1998). Epstein (1995: 4) urges realistic expectations, reminding us not to “let the perfect be the enemy of the good.” Many recommend using PM to create a shift towards a culture of improvement (Proctor and Campbell 1999; Bishop and Pelletier 2001; McLoughlin et al. 2001). In the United States, Braun et al. (1999) and others suggest a national, staged approach including standardized core measures. Berwick (1998) presents an insightful review that challenges current assumptions about health-care performance. Finally, Lied and Sheingold (2001: 394) summarize the current state of practice on PM as follows: “There are real concerns that the act of measurement itself has taken on such a symbolic significance over and above the power of such information to promote beneficial and worthwhile change. We do not yet know how to make such systems deliver on the promises made for them.”

Finally, there are some key structural aspects of healthcare that challenge actionability. The long and strong tradition of professional autonomy, particularly among physicians, focuses philosophically on individuals, not systems. In many jurisdictions, healthcare professionals have contractual (not employee) relationships with service organizations. There are ethical obligations, real or perceived, to provide often heroic and expensive care even where the likelihood of a successful outcome is small. Optimizing performance in such an environment is different from eliminating inefficiencies in a manufacturing process. Clinical care frequently involves trial and error, particularly where cases are intractably difficult or where the science is imprecise, and what one observer would describe as wasteful, another might view as creative and responsive. These caveats suggest that we pay particular attention to the literature that counsels a balanced, nuanced and comprehensive approach to PM and its uses.

Conclusion

The research literature on PM is expanding daily and the ideas are advancing, but our team has read nothing since completion of the major report that stands out in contradiction with the overall findings presented here. A number of encouraging developments are noted on the policy front in Canada since the review: a recognition of the need for leadership in the federal/provincial/territorial accords on indicator reporting and subsequent comparative national reports, the addition (to Saskatchewan’s Health Quality Council) of three more provincial HQCs (Ontario, Quebec and Alberta) and the establishment of the Canadian Patient Safety Institute. At the same time, the controversial Maclean’s Health Report has come and gone. Much of the current energy is focused on wait times and patient safety. We need to address PM more comprehensively, and work remains as well at the service level – in regions and on the front line. Just as it is no longer acceptable to disseminate clinical treatment without evidence, the stakes are too high to implement healthcare PM without developing the evidence base.

TABLE 4. Points of consensus and divergence in the PM literature

Consensus

- Performance can be measured and improved, and performance measurement can be beneficial
- Performance measures should include non-financial measures with a focus on quality, customer needs and, more broadly, stakeholder needs
- There is a need to move towards more meaningful and strategic measures
- There is a need to dedicate sufficient effort at the conceptualization stage, including consideration of the relevance of proposed measures to system change as well as their potential adverse effects
- PM is a complex and technically challenging exercise that needs appropriate expertise, resource allocation, an evidence base and awareness of the pitfalls
- PM system implementation represents significant organizational change, not just the collection and reporting of data
- More emphasis and effort are needed on “actioning” results for improvement

Divergence

- The extent to which PM systems should be integrated across all levels of an organization and, specifically, whether measurement of management performance and measurement of clinical performance should be integrated processes
- The degree to which measures should change over time or remain static for historical comparison
- The optimal horizontal scope of measures
- The relative emphasis on process vs. outcome measures
- In the health literature, whether or not patient-level outcomes should be measured routinely by clinicians for all patients vs. using sampling or case-based approaches
- The extent to which performance results should be reported publicly
- The extent to which measures have specific utility for consumers and the general public
- The utility and relevance of administrative data for, in particular, outcomes measurement
- The extent of customization vs. standardization of measures

ACKNOWLEDGMENTS

The State of the Science Review was funded by the Alberta Heritage Foundation for Medical Research, and significant in-kind support was received from the Alberta Mental Health Board. Thanks are due to K. Omelchuk, H. Gardiner, S. Newman, S. Clelland, A. Beckie, K. Lewis-Ng, I. Frank, J. Osborne, D. Ma, X. Kostaras and O. Berze for their assistance on parts of the broader review. T. Sheldon and C. Baker provided methodologic consultation, and E. Goldner and S. Lewis reviewed the main report. Findings have been presented in part at Academy Health, Nashville, Tennessee, June 2003; World Psychiatric Association, Paris, France, July 2003; International Conference on the Scientific Basis of Health Services, Washington, DC, September 2003; and American Evaluation Association, Reno, Nevada, November 2003.

Correspondence may be directed to: Carol E. Adair, MSc, PhD, Associate Professor, Depts. of Psychiatry and Community Health Sciences, Room 124, Heritage Medical Research Building, 3330 Hospital Dr. NW, Calgary, Alberta T2N 4N1, Tel: 403-210-8805, Fax: 403-944-3144.

REFERENCES

- Adair, C., L. Simpson, J.M. Birdsell, K. Omelchuk, A. Casebeer, H.P. Gardiner, S. Newman, A. Beckie, S. Clelland, K.A. Hayden and P. Beausejour. 2003. *Performance Measurement Systems in Health and Mental Health Services: Models, Practices and Effectiveness. A State of the Science Review*. Calgary: University of Calgary.
- Agency for Healthcare Research and Quality (AHRQ). 2002. "Understanding Quality Measurement." *Child Health Care Quality Toolbox*. Retrieved March 26, 2006. <<http://www.ahrq.gov/chtoolbox/understsn.htm>>.
- Aller, K. 1996. "Information Systems for the Outcomes Movement." *Healthcare Information Management* 10(1): 37–52.
- Baker, G. and G. Pink. 1995. "A Balanced Scorecard for Canadian Hospitals." *Healthcare Management Forum* 8(4): 7–21.
- Baker, S. 1995. "Use of Performance Indicators for General Practice." *British Medical Journal* 311: 209–10.
- Barnsley, J., L. Lemieux-Charles and G. Baker. 1996. "Selecting Clinical Outcome Indicators for Monitoring Quality of Care." *Healthcare Management Forum* 9(1): 5–21.
- Barrell, J. 2000. "Apples to Apples: The Complexities of Health Care Outcomes Reporting." *Infusion* 6(7): 15–24.
- Baughan, P., C. Armistead and D. Parker. 2002. "Managerial Reflections on the Deployment of Balanced Score Cards." In A. Neely, A. Walters and R. Austin, eds., *Performance Measurement and Management: Research and Action*. Boston: Center for Business Performance, Cranfield University.
- Berwick, D. 1998. "Crossing the Boundary: Changing Mental Models in the Service of Improvement." *International Journal for Quality in Health Care* 10(5): 435–41.
- Bishop, W. and L. Pelletier. 2001. "Interview with a Quality Leader: Janet Corrigan on the Institute of Medicine and Healthcare Quality." *Journal for Healthcare Quality* 23(5): 21–24.
- Bodenheimer, T. 1999. "The American Health Care System: The Movement for Improved Quality in Health Care." *New England Journal of Medicine* 340(6): 488–92.
- Bourne, M., J. Mills, M. Wilcox, A. Neely and K. Platts. 2000. "Designing, Implementing and Updating Performance Measurement Systems." *International Journal of Operations and Production Management* 20(7): 754–71.
- Braun, B., R. Koss and J. Loeb. 1999. "Integrating Performance Measure Data into the Joint Commission Accreditation Process." *Evaluation and the Health Professions* 22(3): 283–97.
- Braun, B. and F. Zibrat. 1996. "Developing an Outcomes Measurement System: The Value of Testing." *American Journal of Medical Quality* 11(2): 57–67.
- Brignall, S. 2002. "The Unbalanced Scorecard: A Social and Environmental Critique." In A. Neely, A. Walters and R. Austin, eds., *Performance Measurement and Management: Research and Action*. Boston: Center for Business Performance, Cranfield University.

- Brook, R., E. McGlynn and P. Shekelle. 2000. "Defining and Measuring Quality of Care: A Perspective from US Researchers." *International Society for Quality in Health Care* 12(4): 281–95.
- Brookfield, D. 1992. "Performance Measurement: Focusing on the Key Issue." *Journal of Management in Medicine* 6(2): 39–45.
- Brown, M. 2002. "Change and Stability in the Canadian Healthcare System." *Expert Reviews of Pharmacoeconomics Outcomes Research* 2(4): 309–12.
- Brownell, M., N. Roos and L. Roos. 2001. "Monitoring Health Reform: A Report Card Approach." *Social Science and Medicine* 52(5): 657–70.
- Camp, R. and A. Tweet. 1994. "Benchmarking Applied to Health Care." *Joint Commission Journal on Quality Improvement* 20(5): 229–38.
- Campbell, S., M. Roland and B. Leese. 2001. "Progress in Clinical Governance: Findings from the First NPCRDC National Tracker Survey of Primary Care Groups/Trusts." *British Journal of Clinical Governance* 6(2): 90–93.
- Collopy, B. 1998. "Health-Care Performance Measurement Systems and the ACHS Care Evaluation Program." *Journal of Quality in Clinical Practice* 18(3): 171–76.
- Corso, L., P. Wiesner, P. Halverson and K. Brown. 2000. "Using the Essential Services as a Foundation for Performance Measurement and Assessment of Local Public Health Systems." *Journal of Public Health Management and Practice* 6(5): 1–18.
- DeBaylo, P. 1999. "Ten Reasons Why the Baldrige Model Works." *Journal for Quality and Participation* 22(1): 24–28.
- DeRosario, J. 1999. "Healthcare System Performance Indicators: A New Beginning for a Reformed Canadian Healthcare System." *Journal for Healthcare Quality* 21(1): 37–41.
- Eddy, D. 1998. "Performance Measurement: Problems and Solutions." *Health Affairs* 17(4): 7–25.
- Elkan, R. and J. Robinson. 1998. "The Use of Targets to Improve the Performance of Health Care Providers: A Discussion of Government Policy." *British Journal of General Practice* 48: 1515–18.
- Epstein, A. 1995. "Performance Reports on Quality – Prototypes, Problems and Prospects." *New England Journal of Medicine* 331(1): 57–61.
- Epstein, M. and J. Manzoni. 1998. "Implementing Corporate Strategy: From Tableaux de Bord to Balanced Scorecards." *European Management Journal* 16(2): 190–203.
- Evans, D., T. Edejer, J. Lauer, J. Frenk and C. Murray. 2001. "Measuring Quality: From the System to the Provider." *International Society for Quality in Health Care* 13(6): 439–46.
- Fawcett, S. and M. Cooper. 1998. "Logistics Performance Measurement and Customer Success." *Industrial Marketing Management* 27(4): 341–57.
- Ford, R., S. Bach and M. Fottler. 1997. "Methods of Measuring Patient Satisfaction in Health Care Organizations." *Health Care Management Review* 22(2): 74–89.
- Goddard, M., R. Mannion and P. Smith. 1998. "Performance Indicators. All Quiet on the Front Line." *Health Service Journal* 108: 24–26.
- Goddard, M., R. Mannion and P. Smith. 1999. "Assessing the Performance of NHS Hospital Trusts: The Role of 'Hard' and 'Soft' Information." *Health Policy* 48(2): 119.
- Goddard, M., R. Mannion and P. Smith. 2000. "Enhancing Performance in Health Care: A Theoretical Perspective on Agency and the Role of Information." *Health Economics* 9(2): 95–107.

- Greenhalgh, J., A. Long, A. Brettle and M. Grant. 1996. "The Value of an Outcomes Information Resource. An Evaluation of the UK Clearing House on Health." *Journal of Management Medicine* 10(5): 55–65.
- Gross, P., B. Braun, S. Kritchevsky and B. Simmons. 2000. "Comparison of Clinical Indicators for Performance Measurement of Health Care Quality: A Cautionary Note." *British Journal of Clinical Governance* 5(4): 202–11.
- Guiffrida, A., T. Gosden, F. Forland et al. 2000. "Target Payments in Primary Care: Effects on Professional Practice and Health Care Outcomes." *Cochrane Database of Systematic Reviews* (3): CD000531.
- Hall, J. 1996. "The Challenge of Health Outcomes." *Journal of Quality in Clinical Practice* 16(1): 5–15.
- Handler, A., M. Issel and B. Turnock. 2001. "A Conceptual Framework to Measure Performance of the Public Health System." *American Journal of Public Health* 91(8): 1235–39.
- Hermann, R., H. Leff, R. Palmer, D. Yang, T. Teller, S. Provost, C. Jakubiak and J. Chan. 2000. "Quality Measures for Mental Health Care: Results from a National Inventory." *Medical Care Research and Review* 57(Suppl. 2): 136–54.
- Hoelzer, S., W. Waechter, A. Stewart, L. Raymond and R. Schweiger. 2001. "Towards Case-Based Performance Measures: Uncovering Deficiencies in Applied Medical Care." *Journal of Evaluation in Clinical Practice* 7(4): 355–63.
- Hoey, J., A. Todkill and K. Flegel. 2002. "What's in a Name? Reporting Data from Public Institutions." *Canadian Medical Association Journal* 166(2): 193–94.
- Holloway, J. 2001. "Investigating the Impact of Performance Measurement." *International Journal of Business Performance Management* 3(2–4): 167–80.
- Ibrahim, J. 2001. "Performance Indicators from All Perspectives." *International Journal for Quality in Health Care* 13(6): 431–32.
- Iezzoni, L. 1997. "The Risks of Risk Adjustment." *Journal of the American Medical Association* 278(19): 1600–7.
- Inamdar, N., R. Kaplan, M. Bower and K. Reynolds. 2002. "Applying the Balanced Scorecard in Healthcare Provider Organizations." *Journal of Healthcare Management* 47(3): 179–96.
- Institute for Healthcare Improvement. 2002. Retrieved March 26, 2006. <<http://www.ihl.org/ihl>>.
- Jarvi, K., R. Sultan, A. Lee, F. Lussing and R. Bhat. 2002. "Multi-Professional Mortality Review: Supporting a Culture of Teamwork in the Absence of Error Finding and Blame-Placing." *Hospital Quarterly* 5(4): 58–61.
- Jencks, S. 2000. "Clinical Performance Measurement – A Hard Sell." *Journal of the American Medical Association* 283(15): 2015–16.
- Jennings, B. and N. Staggers. 1999. "A Provocative Look at Performance Measurement." *Nursing Administration Quarterly* 24(1): 17–30.
- Johnston, R. and L. Fitzgerald. 2001. "Performance Measurement: Flying in the Face of Fashion." *International Journal of Business Performance Management* 3(2–4): 181–90.
- Kanji, G. and P. Moura. 2002. "Kanji's Business Scorecard." *Total Quality Management* 13(1): 13–27.

- Kaplan, R. and D. Norton. 1996. "Linking the Balanced Scorecard to Strategy." *California Management Review* 39(1): 53–79.
- Kaplan, R. and D. Norton. 2001. "Transforming the Balanced Scorecard from Performance Measurement to Strategic Management: Part I." *Accounting Horizons* 15(1): 87–104.
- Kates, J., K. Marconi and T. Mannle Jr. 2001. "Developing a Performance Management System for a Federal Public Health Program: The Ryan White CARE Act, Titles I and II." *Evaluation Program and Planning* 24(2): 145–55.
- Kelman, C. and L. Smith. 2000. "It's Time: Record Linkage – The Vision and the Reality." *Australian and New Zealand Journal of Public Health* 24(1): 100–1.
- Kizer, K. 2001. "Establishing Health Care Performance Standards in an Era of Consumerism." *Journal of the American Medical Association* 286(10): 1213–17.
- Klazinga, N., K. Stronks, D. Delnoij and A. Verhoeff. 2001. "Indicators without a Cause. Reflections on the Development and Use of Indicators in Health Care from a Public Health Perspective." *International Journal for Quality in Health Care* 13(6): 433–38.
- Kleinpell, R. 1997. "Whose Outcomes: Patients, Providers, or Payers?" *Nursing Clinics of North America* 32(3): 513–20.
- Kueng, P. 2000. "Process Performance Measurement System: A Tool to Support Process-Based Organizations." *Total Quality Management* 11(1): 67–85.
- Kueng, P. and A. Krahn. 1999. "Building a Process Performance Measurement System: Some Early Experiences." *Journal of Scientific and Industrial Research* 58(3–4): 149–59.
- Leatherman, S. and D. McCarthy. 1999. "Public Disclosure of Health Care Performance Report." *International Journal for Quality in Health Care* 11(2): 93–105.
- Lebas, M. 1995. "Performance Measurement and Performance Management." *International Journal of Production Economics* 41(1–3): 23–35.
- Leggat, S., L. Narine, L. Lemieux-Charles, J. Barnsley, G. Baker, C. Sicotte, F. Champagne and H. Bilodeau. 1998. "A Review of Organizational Performance Assessment in Health Care." *Health Services Management Research* 11: 3–23.
- Legnini, M., L. Rosenberg, M. Perry and N. Robertson. 2000. "Where Does Performance Measurement Go from Here?" *Health Affairs* 19(3): 173–77.
- Lemieux-Charles, L., N. Gault, F. Champagne, J. Barnsley, I. Trabut, C. Sicotte and D. Zitner. 2000. "Use of Mid-Level Indicators in Determining Organizational Performance." *Hospital Quarterly* 3(4): 48–52.
- Lemieux-Charles, L., W. McGuire, F. Champagne, J. Barnsley, D. Cole and C. Sicotte. 2002. "Multilevel Performance Indicators: Examining Their Use in Managing Performance in Health Care Organizations." In A. Neely, A. Walters and R. Austin, eds., *Performance Measurement and Management: Research and Action*. Boston: Center for Business Performance, Cranfield University.
- Lied, T. 1999. "Performance: A Multi-Disciplinary and Conceptual Model." *Journal of Evaluation in Clinical Practice* 5(4): 393–400.
- Lied, T. and S. Sheingold. 2001. "Relationships among Performance Measures for Medicare Managed Care Plans." *Health Care Financing Review* 22(3): 23–33.
- Lockamy III, A. 1998. "Quality-Focused Performance Measurement Systems: A Normative

- Model." *International Journal of Operations and Production Management* 18(8): 740–66.
- Luttman, R. 1998. "Next Generation Quality, Part 2: Balanced Scorecards and Organizational Improvement." *Topics in Health Information Management* 19(2): 22–29.
- Mannion, R. and H. Davies. 2002. "Reporting Health Care Performance: Learning from the Past, Prospects for the Future." *Journal of Evaluation in Clinical Practice* 8(2): 215–28.
- Mant, J. and N. Hicks. 1996. "Assessing Quality of Care: What Are the Implications of the Potential Lack of Sensitivity of Outcome Measures to Differences in Quality?" *Journal of Evaluation in Clinical Practice* 2(4): 243–48.
- Mark, B., J. Salyer and N. Geddes. 1997. "Outcomes Research. Clues to Quality and Organizational Effectiveness?" *Nursing Clinics of North America* 32(3): 589–601.
- Marshall, M., P. Shekelle, S. Leatherman and R. Brook. 2000. "The Public Release of Performance Data: What Do We Expect to Gain? A Review of the Evidence." *Journal of the American Medical Association* 283(14): 1866–74.
- McCormick, D., D. Himmelstein, S. Woolhandler, S. Wolfe and D. Bor. 2002. "Relationship between Low Quality-of-Care Scores and HMOs' Subsequent Public Disclosure of Quality-of-Care Scores." *Journal of the American Medical Association* 288(12): 1484–90.
- McEwan, K. and E. Goldner. 2000. *Accountability and Performance Indicators for Mental Health Services and Supports*. Prepared for the Federal/Provincial/Territorial Advisory Network on Mental Health. Ottawa: Health Canada.
- McIntyre, D., L. Rogers and E. Heier. 2001. "Overview, History and Objectives of Performance Measurement." *Health Care Financing Review* 22(3): 7–21.
- McKee, M. and P. James. 1997. "Using Routine Data to Evaluate Quality of Care in British Hospitals." *Medical Care* 35(10) (Suppl.): OS102–11.
- McKee, M. and T. Sheldon. 1998. "Measuring Performance in the NHS." *British Medical Journal* 316(7128): 322.
- McLoughlin, V., S. Leatherman, M. Fletcher and J. Owen. 2001. "Improving Performance Using Indicators. Recent Experiences in the United States, the United Kingdom, and Australia." *International Journal for Quality in Health Care* 13(6): 455–62.
- Mooraj, S., D. Oyon and D. Hostettler. 1999. "The Balanced Scorecard: A Necessary Good or an Unnecessary Evil?" *European Management Journal* 17(3): 481–91.
- Morgan, C. and A. Braganza. 2002. "Performance Measurement Systems: Knowledge Developer or Destroyer?" In A. Neely, A. Walters and R. Austin, eds., *Performance Measurement and Management: Research and Action*. Boston: Center for Business Performance, Cranfield University.
- Moscovice, I., J. Christianson and A. Wellever. 1995. "Measuring and Evaluating the Performance of Vertically Integrated Rural Health Networks." *Journal of Rural Health* 11(1): 9–21.
- Nadzam, D. and M. Nelson. 1997. "The Benefits of Continuous Performance Measurement." *Nursing Clinics of North America* 32(3): 543–59.
- Naylor, G. 1999. "Using the Business Excellence Model to Develop a Strategy for Healthcare Organisation." *International Journal of Health Care Quality Assurance* 12(2): 37–44.
- Neely, A. 1999. "The Performance Measurement Revolution: Why Now and What Next?" *International Journal of Operations and Production Management* 19(2): 205–28.

- Neely, A., M. Gregory and K. Platts. 1995. "Performance Measurement System Design – A Literature Review and Research Agenda." *International Journal of Operations and Production Management* 15(4): 80–116.
- Neely, A., J. Mills, K. Platts, H. Richards, M. Gregory, M. Bourne and M. Kennerley. 2000. "Performance Measurement System Design: Developing and Testing a Process-Based Approach." *International Journal of Operations and Production Management* 20(9–10): 1119–45.
- Newhouse, J. 2002. "Why Is There a Quality Chasm?" *Health Affairs* 21(4): 13–25.
- Nutley, S. and P. Smith. 1998. "League Tables for Performance Improvement in Health Care." *Journal of Health Services and Research Policy* 3(1): 50–57.
- Pink, G., I. McKillop, E. Schraa, C. Preyra, C. Montgomery and G. Baker. 2001. "Creating a Balanced Scorecard for a Hospital System." *Journal of Health Care Finance* 27(3): 1–20.
- Proctor, S. and C. Campbell. 1999. "A Developmental Performance Framework for Primary Care." *International Journal of Health Care Quality Assurance* 12(7): 279–86.
- Roper, W. and G. Mays. 2000. "Performance Measurement in Public Health: Conceptual and Methodological Issues in Building the Science Base." *Journal of Public Health Management and Practice* 6(5): 66–77.
- Rubin, H., P. Pronovost and G. Diette. 2001. "The Advantages and Disadvantages of Process-Based Measures of Health Care Quality." *International Journal for Quality in Health Care* 13(6): 469–74.
- Schneider, E. 2002. "Measuring Mortality Outcomes to Improve Health Care: Rational Use of Ratings and Rankings." *Medical Care* 40(1): 1–3.
- Schneider, E., V. Riehl, S. Courte-Wienecke, D. Eddy and C. Sennett. 1999. "Enhancing Performance Measurement: NCQA's Road Map for a Health Information Framework." *Journal of the American Medical Journal* 282(12): 1184–90.
- Sennett, C. 1998. "Moving Ahead, Measure by Measure." *Health Affairs* 17(4): 36–38.
- Shahian, D., S. Normand, D. Torchiana, S. Lewis, J. Pastore, R. Kuntz and P. Dreyer. 2001. "Cardiac Surgery Report Cards: Comprehensive Review and Statistical Critique." *Annals of Thoracic Surgery* 72(6): 2155–68.
- Shaw, C. 1997. "Health-Care League Tables in the United Kingdom." *Journal of Quality in Clinical Practice* 17(4): 215–19.
- Sheldon, T. 1994. "Please Bypass the PORT: Observational Studies of Effectiveness Run a Poor Second to Randomized Controlled Trials." *British Medical Journal* 309(6948): 142–43.
- Sheldon, T. 1998. "Promoting Health Care Quality: What Role Performance Indicators?" *Quality in Health Care* 7(Suppl.): s45–s50.
- Slater, C. 1997. "What Is Outcomes Research and What Can It Tell Us?" *Evaluation and the Health Professions* 20(3): 243–64.
- Smith, P. 1993. "Outcome-Related Performance Indicators and Organizational Control in the Public Sector." *British Journal of Clinical Governance* 4(3): 135–51.
- Smith, P. 2002. "Performance Management in British Health Care: Will It Deliver?" *Health Affairs* 21(3): 103–15.
- Smith, P. and M. Goddard. 2002. "Performance Management and Operational Research: A

- Marriage Made in Heaven?" *Journal of the Operational Research Society* 53(3): 247–55.
- Stryer, D., S. Tunis, H. Hubbard and C. Clancy. 2000. "The Outcomes of Outcomes and Effectiveness Research: Impacts and Lessons from the First Decade." *Health Services Research* 35(5, Part 1): 977–93.
- Tennant, C. and P.A. Roberts. 2000. "A Technique for Strategic Quality Management." *Quality Assurance* 8(2): 77–90.
- Turpin, R., L. Darcy, R. Koss, C. McMahonill, K. Meyne, D. Morton, J. Rodriguez, S. Schmaltz, P. Schyve and P. Smith. 1996. "A Model to Assess the Usefulness of Performance Indicators." *International Journal for Quality in Health Care* 8(4): 321–29.
- Ullman, M., C. Metzger, T. Kuzel and C. Bennett. 1996. "Performance Measurement in Prostate Cancer Care: Beyond Report Cards." *Urology* 47(3): 356–65.
- van Peurse, K., M. Pratt and S. Lawrence. 1995. "Health Management Performance: A Review of Measures and Indicators." *Accounting, Auditing and Accountability Journal* 8(5): 34–70.
- Viccars, A. 1998. "Clinical Governance: Just Another Buzzword of the 90's?" *MIDIRS Midwifery Digest* 8(4): 409–12.
- Voelker, K., J. Rakich and G. French. 2001. "The Balanced Scorecard in Healthcare Organizations: A Performance Measurement and Strategic Planning Method." *Hospital Topics* 79(3): 13–24.
- Waggoner, D., A. Neely and M. Kennerley. 1999. "The Forces That Shape Organisational Performance Measurement Systems: An Interdisciplinary Review." *International Journal of Production Economics* 60–61: 53–60.
- Ward, S. 2000. "Counting on Quality." *Nursing Standard* 14(52): 16.
- Weinberg, N. 2001. "Using Performance Measures to Identify Plans of Action to Improve Care." *Joint Commission Journal on Quality Improvement* 27(12): 683–88.
- West, R. 1996. "NHS Performance Guides: Raising the Standard – Indirectly?" *Journal of Public Health Medicine* 19(3): 361–63.
- Williams, I., D. Naylor, M. Cohen, V. Goel, A. Basinski, L. Ferris and H. Llewellyn-Thomas. 1992. "Outcomes and the Management of Health Care." *Canadian Medical Association Journal* 147(12): 1775–80.
- Wynia, M., R. Hasnain-Wynia, E. McGlynn and R. Brook. 1996. "Assessing Quality of Care: Process Measures vs. Outcomes Measures." *Journal of the American Medical Association* 276(19): 1551–52.
- Zairi, M. and Y. Jarrar. 2001. "Measuring Organizational Effectiveness in the NHS: Management Style and Structure Best Practices." *Total Quality Management* 12(7, 8): 882–89.
- Zaslavsky, A. 2001. "Statistical Issues in Reporting Quality Data: Small Samples and Casemix Variation." *International Journal for Quality in Health Care* 13(6): 481–88.